Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **486,** 582–595 (2019)
Advance Access publication 2019 March 20

doi:10.1093/mnras/stz824

The effect of assembly bias on redshift-space distortions

N. Padilla,[1,2]★ S. Contreras,[1,3,4] I. Zehavi,[5] C. M. Baugh[6] and P. Norberg[6,7]

[1]*Institúto Astrofísica, Pontifica Universidad Católica de Chile, Santiago 782-0436, Chile*
[2]*Centro de Astro-Ingeniería, Pontificia Universidad Católica de Chile, Santiago 782-0436, Chile*
[3]*Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Plaza San Juan 1, Planta-2, Teruel E-44001, Spain*
[4]*Donostia International Physics Center (DIPC), Manuel Lardizabal pasealekua 4, E-20018 Donostia, Basque Country, Spain*
[5]*Department of Physics, Case Western Reserve University, Cleveland, OH 44106, USA*
[6]*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[7]*Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

Accepted 2019 March 18. Received 2019 March 14; in original form 2018 September 15

ABSTRACT

We study potential systematic effects of assembly bias on cosmological parameter constraints from redshift-space distortion measurements. We use a semi-analytic galaxy formation model applied to the Millennium *N*-body *WMAP*-7 simulation to study the effects of halo assembly bias on the redshift-space distortions of the galaxy correlation function. We look at the pairwise velocities of galaxies living in haloes with concentrations and ages in the upper and lower quintiles, and find that the velocity differences between these are consistent with those reported for real-space clustering analyses; i.e. samples with higher clustering also exhibit stronger infall pairwise motions. This can also be seen in the monopole and quadrupole of the redshift-space correlation function. We find that regardless of the method of measurement, the changes in the β parameter due to different secondary halo parameters fully track the change in the bias parameter. Hence, assembly bias does not introduce detectable systematics in the inferred logarithmic growth factor.

Key words: galaxies: evolution – galaxies: formation – galaxies: haloes – galaxies: statistics – large-scale structure of Universe – cosmology: theory.

1 INTRODUCTION

One of the most successful tools to determine the cosmological parameters of the Universe is the analysis of the galaxy correlation function measured in redshift space (Hu & Haiman 2003; Guzzo et al. 2008; Wagner, Müller & Steinmetz 2008; Shoji, Jeong & Komatsu 2009). This correlation function differs from the real-space correlation function since the inferred comoving positions of galaxies are shifted by their peculiar velocities in the direction of the line of sight, which introduces an anisotropy in the correlation function. Its strength depends on the amplitude of peculiar velocities, which in turn encodes cosmological information; the rate and acceleration of the expansion of the Universe oppose gravitational collapse and this is encoded in the rate of growth of perturbations that is revealed by the peculiar motions of galaxies (Peebles 1980).

In some cases the full two-dimensional correlation function $\xi(\mu, s)$ is used to make constraints on cosmological parameters, where the dimensions correspond to the pair separation in redshift space, s, and the cosine of the angle to the line of sight, μ (Okumura et al. 2008; Chuang & Wang 2012). However, the covariance matrices needed to estimate errors in the parameters are too large owing to the large number of bins of this two-dimensional space. This makes the covariance matrix subject to important uncertainties. Padmanabhan & White (2008) proposed to use one-dimensional multipole functions obtained using Legendre polynomials to alleviate this problem and, at the same time, to avoid assuming knowledge of the shape of redshift-space distortions; they show that the joint analysis of the monopole and quadrupole provides measurements of the angular diameter distance $D_A(z)$ and the Hubble constant $H(z)$ (see also the 'clustering wedges' method proposed by Kazin, Sánchez & Blanton 2012 that constrains the same quantities).

When inferring cosmological parameters from redshift-space distortions, it is commonly assumed that galaxy velocities are unbiased with respect to those of the underlying matter. On small scales there is evidence that this is not necessarily a good approximation. Guo et al. (2015) studied the CMASS BOSS sample and found that central galaxies show a velocity dispersion of ∼20 per cent of the total velocity dispersion of the haloes, and that the velocity dispersion of satellites is ∼80 per cent of the total. Furthermore, Ye et al. (2017) reach similar conclusions for the SDSS Main sample analysing jointly with galaxies from the Illustris simulation. On slightly larger scales, Hearin (2015) shows that the halo assembly history, in addition to halo mass, affects the galaxy pairwise velocity dispersion by about 20 per cent in their amplitude up to scales of $10\,h^{-1}$ Mpc, implying that models of the galaxy distribution that

★ E-mail: npadilla@astro.puc.cl

intend to use the dynamics of galaxies on these scales need to take this effect into account. On scales corresponding to galaxies hosted by different haloes (the two-halo term), it is generally assumed that galaxy pairwise velocities are unbiased. This appears to be a good approximation; a study of the velocity bias b_v calculated as the square root of the ratio between velocity power spectra of haloes and that of the mass in a cosmological numerical simulation by Chen et al. (2018) shows that there are significant deviations from $b_v = 1$ that depend on redshift and mass, but these are of the order of 1 per cent out to $k = 0.2\,h\,\mathrm{Mpc}^{-1}$ at $z = 0$, and can reach 5 per cent on the same scales at higher redshifts, $z \sim 2$.

Halo clustering was initially assumed to depend on halo mass alone. However, Gao, Springel & White (2005) showed that it also depends on halo assembly history (this was also noticed in Sheth & Tormen 2004). Croton, Gao & White (2007) termed this effect 'assembly bias' and demonstrated that it also impacts galaxy clustering. The definition of halo assembly bias was expanded to encapsulate the dependence of (large-scale) clustering amplitude on other halo properties in addition to halo mass, i.e. a second parameter such as spin, concentration, or the number of substructures per halo (see e.g. Gao & White 2007; Mao, Zentner & Wechsler 2018; Salcedo et al. 2018; Sato-Polito et al. 2018 for recent compilations of properties that produce halo assembly bias).

Occupancy variation is the dependence of the halo occupation distribution (i.e. the number of galaxies per halo as a function of halo mass; HOD from this point on) on halo properties in addition to mass (Zehavi et al. 2018). This in turn can induce changes in the observed clustering of the galaxies that populate these haloes (Zehavi et al. 2018; see also Artale et al. 2018; Contreras et al. 2019). The combined effect of halo assembly bias and occupancy variation determines the amplitude of clustering of a given sample of galaxies. When both are present, such a sample is said to be affected by galaxy assembly bias; i.e. these galaxies show a dependence of their clustering amplitude on halo properties other than their mass.

Despite several attempts to find hard evidence of assembly bias in galaxy surveys, it is still not clear whether the different signatures pointed out by various authors are due to systematic effects that can mimic the effects of assembly bias (e.g. Campbell et al. 2015; Zu & Mandelbaum 2016; Busch & White 2017; Sin, Lilly & Henriques 2017; Lacerna et al. 2018; Tinker et al. 2018). To date, not only is it still not proven that galaxy assembly bias is present in the real Universe, it is also not clear if it would have an impact on cosmological constraints, as shown, for instance, by McEwen & Weinberg (2018). They show that the cross-correlation coefficient between galaxies and mass shows no effect of assembly bias on scales of 1–10 $h\,\mathrm{Mpc}^{-1}$, and on larger scales the cross-correlation is expected to be unity in any case. Cosmological constraints from the combination of galaxy–galaxy lensing and galaxy clustering should therefore be unaffected by galaxy assembly bias, even if the galaxy autocorrelation and galaxy-mass cross-correlation functions are affected individually.

The relation between pairwise velocities and the integrated overdensity $\Delta(r)$ reads

$$v(r) = -\frac{1}{3}\frac{H(z)}{(1+z)}\frac{f(\Omega)}{b} r\Delta(r), \quad (1)$$

where b is the bias factor and $f(\Omega)$ is the logarithmic growth rate, which depends on the matter density parameter Ω (Peebles 1980). If pairwise velocities are indeed unbiased in the two-halo-term regime, then the assembly bias effect seen on spatial clustering statistics should translate completely to phase-space statistics simply as a different bias factor and thus not introduce any systematic effects in cosmological parameter constraints. This is what we intend to look into here. To do so we use semi-analytic galaxies from a numerical simulation and find out whether velocity differences of galaxies living in haloes of equal mass but with different secondary parameters can be completely explained by differences in their bias parameter as inferred from spatial statistics (i.e. in real space).

Our work complements recent results by Xu & Zheng (2018), who look at the change of pairwise velocities of dark matter haloes as a function of mass and several secondary parameters. They find that the changes in pairwise velocities and clustering due to secondary parameters are correlated but that the scatter in this relation is wide. Here we focus on galaxy populations from a semi-analytic model (SAM) and look at the effect of halo concentration and age. In addition we will also look at the effect of galaxy selection and evolution with redshift. Finally, rather than looking at the correlation between changes in velocity and clustering, we will focus on whether cosmological parameters obtained from redshift-space distortions could be biased because of galaxy assembly bias.

This work is the latest installment in a series of papers focusing on galaxy assembly bias. In Zehavi et al. (2018) we made a careful analysis of the variation of the number of galaxies in haloes due to different secondary properties and how this combines with halo assembly bias to change the clustering of galaxies. The next paper in this series was Contreras et al. (2019), where we looked at the evolution of galaxy assembly bias with redshift. In the present installment, we extend the analysis in these two previous works to velocity space, with special emphasis on redshift-space distortions.

This paper is organized as follows. In Section 2 we present the *N*-body simulation, the semi-analytic model that is applied to it, and the sample selection. In Section 3 we present the method to create shuffled samples of galaxies that remove the occupancy variation but retain the inner halo structure in both position and velocity space. Section 4 presents results of the pairwise velocities and their dependence on parameters other than the halo mass. We extend the analysis to correlation functions in redshift space in Section 5, and Section 6 shows the cosmological parameters one would estimate from these correlation functions. Finally we present our conclusions in Section 7. The appendix contains additional tests and different ways to calculate the effect of galaxy assembly bias on the β parameter.

2 SIMULATIONS AND SAMPLES

We use numerical simulations populated with a semi-analytic model of galaxy formation. SAMs of galaxy formation couple the output of large dark-matter-only simulations with simple descriptions of the physical processes affecting the baryonic content of galaxies to help study galaxy formation on a cosmological footing. SAMs follow a galaxy population as it evolves, which allows us to study galaxy samples at different redshifts, selected according to different physical properties. We outline the details of the simulation and the selection of samples below.

2.1 Semi-analytic model and numerical simulation

The typical processes that are followed in semi-analytic models are: the collapse and merging of dark matter haloes, shock heating and radiative cooling of gas, star formation, supernovae, active galactic nuclei, and photoionization feedback, chemical enrichment of the gas and stars, instabilities of the gaseous disc, and galaxy mergers.

We will focus our analysis on the Guo et al. (2013; hereafter G13) model, which is a version of L-GALAXIES, the SAM code of

the Munich group (De Lucia, Kauffmann & White 2004; Croton et al. 2006; De Lucia & Blaizot 2007; Guo et al. 2011; Henriques et al. 2013, 2015). An in-depth description of the model can be found in Guo et al. (2013; see also Guo et al. 2016; Contreras et al. 2017a).

We use outputs obtained after this model is applied to the Millennium *WMAP*7 simulation (Guo et al. 2013), which are publicly available.[1] This simulation shares several characteristics with the Millennium simulation (Springel et al. 2005), but is run using a *WMAP*7 cosmology.[2] The simulation follows 2160^3 particles in a $(500\,h^{-1}$ Mpc$)^3$ periodic comoving box, with a particle mass of $9.31 \times 10^8\,h^{-1}\,\mathrm{M}_\odot$. In total, 61 simulation outputs were stored between $z = 50$ and $z = 0$.

The simulation outputs are used to construct merger trees, which are an input to the SAM. The haloes are identified using a friends-of-friends (FoF) finding algorithm (Davis et al. 1985). This is run on each output of the simulation and all haloes with at least 20 particles are saved. SUBFIND is then run on these halo catalogues to identify their substructures, called subhaloes (Springel et al. 2001). The final step is to use these subhaloes to build the merger trees by linking each subhalo in one snapshot to a single descendant subhalo in the following output. The mass of the halo ($M_{\rm halo}$) is defined as the mass within the radius where the halo overdensity is 200 times the critical density (referred to as 'm_crit200' in the public data base).

Galaxies in the G13 model are assigned velocities following the approach set out in Guo et al. (2011). While galaxies are associated with a dark matter subhalo (either central or satellite) they are assigned the subhalo velocity until the baryonic mass of the galaxy becomes larger than the dark matter mass of the subhalo. This happens to subhaloes due to tidal stripping. This process has a stronger effect on extended dark matter haloes compared to the more concentrated baryons in galaxies, which are also subject to stripping processes followed analytically by the model. Because the orbit of the subhalo is not accurate after this limit, it is replaced with that of the most bound particle of the subhalo, modified by a time-dependent orbit-shrinking factor that models the orbital decay caused by the dynamical friction. This treatment of satellite velocities produces density profiles of clusters of galaxies in good agreement with observations, but it is analytic and subject to several assumptions. We will bear this in mind when analysing the redshift-space distortions measured from the model output.

2.2 Classifying by halo properties

Assembly bias in haloes appears in cosmological simulations when studying samples of equal mass and different secondary property. We will concentrate on two different secondary properties, the halo formation time, with which we separate our galaxies into those living in late- and early-formed haloes (low and high formation redshift), and the halo concentration to define low- and high-concentration samples.

We adopt the redshift at which the halo reaches half its final mass as its formation time, interpolating between the available simulation snapshots (e.g. Gao et al. 2004; Gao et al. 2005; Croton et al. 2007; Zehavi et al. 2018). As it evolves, a halo sometimes shows spurious changes of mass (Contreras, Padilla & Lagos 2017b); to avoid being affected by this problem we set the formation time as the first occurrence of a halo reaching half its final mass. We calculate the formation redshift at each snapshot; therefore, it is expected that as a halo evolves, its formation time defined in this way also changes.

Halo concentration is defined using the ratio between $V_{\rm max}$ and $V_{\rm vir}$ (Gao & White 2007), where $V_{\rm max}$ is the peak value of its circular velocity curve ($V_{\rm c} = (G\,M(r)/r)^{1/2}$) and $V_{\rm vir}$ is the virial velocity of a halo ($V_{\rm vir} \equiv V_{\rm c}(r_{\rm vir})$) (Bullock et al. 2001; Gao & White 2007). This definition has the advantage that it can be calculated directly with the information available in the public data base without the need for model fitting. The halo concentration can show a strong halo assembly bias effect and is advantageous with respect to age as it can be obtained from single snapshots rather than merger trees (e.g. Gao et al. 2005; Wechsler et al. 2006,Contreras et al. 2019). In addition, there are methods to estimate halo concentration in observations (e.g. Biviano et al. 2017). Because of this we will pay special attention to the results using this secondary parameter.

To define our samples of early- and late-formation haloes (and of high and low concentration) we take the upper and lower quintile of the samples in bins of halo mass of 0.1 dex, effectively forcing the mass function to be the same in each sample. We test whether the bin size affects studies of assembly bias in Zehavi et al. (2018) and find little changes for formation time in $z = 0$ results. Contreras et al. (2019) show that halo age is reliably measured in the Millennium *WMAP*-7 simulation up to $z = 3$.

2.3 Galaxy selection

We use the number density of galaxies ranked in order of decreasing stellar mass and star formation rate (SFR) to construct galaxy samples from the G13 model. We will use the following number densities, $n = 0.0316, 0.01, 0.00316, 0.001$, and $0.000316\,h^3\,\mathrm{Mpc}^{-3}$ and redshifts $z = 0$, 1, 2, and 3. It is quite common in the literature to use number density to define samples of galaxies from surveys, and in cases where passive evolution can be safely assumed (eg. Seo, Eisenstein & Zehavi 2008) to connect galaxy populations at different redshifts (Padilla et al. 2010; Leja et al. 2013; Mundy, Conselice & Ownsworth 2015; Torrey et al. 2015; Contreras et al. 2017a). A number density selection has the definite advantage of being insensitive to systematics in the estimation of the galaxy property that is selected to rank the galaxies, and can be applied both to simulations and to observations, which facilitates meaningful comparisons between them.

A stellar mass selection is similar to the selection used in optical or near-infrared surveys, whereas the SFR is relevant to observational samples selected by emission-line strength or rest-frame UV luminosity. The samples were chosen to be evenly spaced in logarithmic number density, corresponding to a change of half a decade in log abundance.

We will place special emphasis on the stellar-mass-selected samples that mimic the SDSS main galaxy sample (Strauss et al. 2002), with $z = 0$ and $n = 0.001\,h^3\,\mathrm{Mpc}^{-3}$, the luminous red galaxies of SDSS-III and IV (BOSS, Eisenstein et al. 2011; and eBOSS, Dawson et al. 2016) with $z \sim 1.0$ and $n = 0.00031\,h^3\,\mathrm{Mpc}^{-3}$, and also on SFR-selected samples similar to the eBOSS emission-line galaxy sample (ELG; Comparat et al. 2016) corresponding to $z \sim 1.0$ and $n = 0.000316\,h^3\,\mathrm{Mpc}^{-3}$, and to HETDEX (Adams et al. 2011) with $z \sim 2.0$ and $n = 0.0001\,h^3\,\mathrm{Mpc}^{-3}$. To illustrate the procedure by which we select our samples we show the cumulative stellar mass and SFR functions in the top and bottom panels of Fig. 1 for all the redshifts studied. The horizontal dashed lines show the number density cuts that will be used throughout. For a given

[1] http://gavo.mpa-garching.mpg.de/Millennium/

[2] The cosmological parameters used in the simulation are $\Omega_{\rm m0} = \Omega_{\rm dm0} + \Omega_{\rm b0} = 0.272$, $\Omega_{\Lambda 0} = 0.728$, $\Omega_{\rm b0} = 0.0455$, $\sigma_8 = 0.81$, $n_{\rm s} = 0.967$, $h = 0.704$.

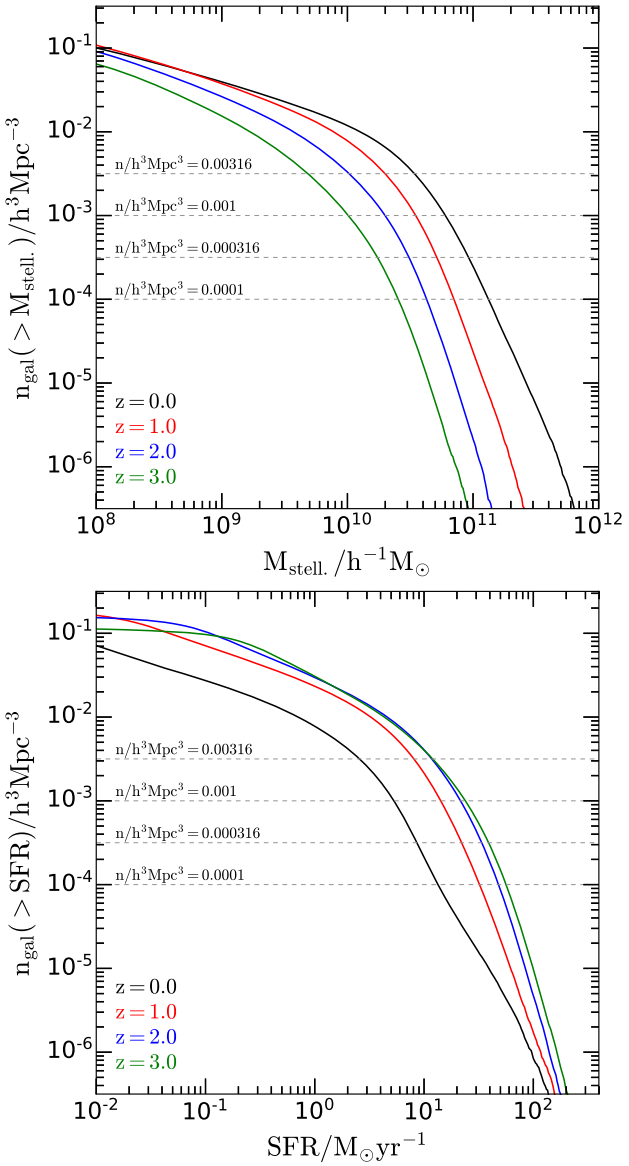

Figure 1. Cumulative stellar mass and SFR functions (top and bottom panels, respectively) at $z = 0$, 1, 2, and 3 (different colours indicated in the key). The dashed horizontal lines represent the number densities used to select samples of galaxies in this work as labelled.

number density, samples of galaxies are constructed using masses or SFRs higher than their value at the intersection with the associated dashed line.

In order to analyse the clustering of galaxies in redshift space, we use the Z-coordinate of the peculiar velocities to displace the Z-axis component of a galaxy position, $Z_s = Z + v_Z/H(z)$, where the subscript s denotes a position in redshift space.

3 SHUFFLED CATALOGUES WITH VELOCITY INFORMATION

Statistics of the phase-space density fields of galaxies, such as galaxy clustering in redshift space, should in principle respond to two effects, halo assembly bias on the one hand and occupancy variation on the other. The halo properties that produce halo assembly bias, i.e. a dependence of halo clustering on secondary properties in addition to their mass, include halo age, spin, number of substructures, and concentration (e.g. Gao & White 2007). The effect of occupancy variation also changes the amplitude of galaxy clustering, as haloes of different secondary properties are populated by different numbers of galaxies; together they give rise to galaxy assembly bias.

The process of shuffling galaxies between haloes in simulations effectively removes the HOD dependence on halo properties; i.e. it removes the occupancy variation component of galaxy assembly bias. This technique was originally proposed by Croton et al. (2007) and consists of interchanging the galaxy population of individual haloes in narrow ranges of halo mass, thus removing the connection to the assembly history of the haloes. This way haloes of similar mass and different properties do not show differences in the number of galaxies that populate them. In this work we adopt a halo mass bin of 0.1 dex in $h^{-1}\,M_\odot$. Central and satellite galaxies in individual haloes are moved together to a different one while retaining their positions relative to the centre of potential of the halo, where the central galaxy of a halo is placed.

An alternative shuffling procedure has been proposed in the literature with the satellites shuffled as well among different haloes of the same mass independently of the central galaxy population (e.g. Zu et al. 2008; Zentner, Hearin & van den Bosch 2014). This shuffling is important if one is concerned with the relation between centrals and satellites, or between satellites in the same halo, such as in the case of satellite–central alignments. In our case this is not needed as we will treat centrals and satellites simply as tracers of the density field.

The comparison of the clustering of shuffled and unshuffled samples helps to quantify whether galaxy assembly bias is present or not. Additionally, one can also divide both samples according to a secondary halo property to learn the role that halo assembly bias and occupancy variation play in setting the amplitude of clustering for this particular property; an example of this approach can be found in the discussion of fig. 10 of Contreras et al. (2019). Namely, if the clustering of the original sample changes with the secondary property, while that of the shuffled one does not, this indicates that occupancy variation is driving the dependence of clustering. In the case where the shuffled samples also show the change in clustering, the reason behind it is halo assembly bias.

Here we extend the shuffling procedure to velocity space, where before being assigned to a new halo, the velocities of halo member galaxies are expressed relative to the velocity of the centre of potential of the haloes. This process retains the relative velocities between galaxies in different haloes such as the Kaiser effect but removes, for instance, the dependence of velocity dispersion on secondary halo parameters, as well as the possible alignments between the velocity and position ellipsoids of satellite systems in neighbouring haloes (Forero-Romero, Contreras & Padilla 2014).

This shuffling method retains the phase-space structure of haloes, and allows us to extend assembly bias studies to velocity space. Namely, we will look into whether assembly bias affects velocities in a different way than positions. Such a difference could lead to potential systematic biases in cosmological parameter constraints obtained from redshift-space distortion analyses.

4 ASSEMBLY BIAS AND VELOCITY STATISTICS

In this section we present measurements of three-dimensional average radial velocities between pairs of galaxies. We refer to this statistic as 'pairwise velocities'; these velocities are the ones

responsible for the distortion of the correlation function in the direction parallel to the line of sight, as they change the relative separation of pairs along this coordinate when galaxy distances are obtained via their redshift (Hamilton 1992). We will measure how the pairwise velocities change for different halo secondary properties, and analyse what effect the assembly bias has on these measurements bearing in mind the possibility that changes in the pairwise velocities simply reflect the overall change in the bias factor of the different subsamples.

4.1 Pairwise velocities

We will show measurements of pairwise velocities between the galaxies in our samples in different concentration (halo age) ranges, and the full sample. This will allow us to maximize the signal to noise of our measurements.

For the pairwise velocities we calculate

$$\boldsymbol{v}_{12}(\boldsymbol{r}) = \left\langle (v_i - v_j)\hat{\boldsymbol{r}_{ij}} \right\rangle |_r, \quad (2)$$

where $\boldsymbol{r} = |r_i - r_j|$, $\hat{\boldsymbol{r}_{ij}} = (r_i - r_j)/\boldsymbol{r}$, i and j refer to a pair of galaxies, $\boldsymbol{v}$ and $\boldsymbol{r}$ are their comoving velocity and position vectors, and the average is done on pairs separated by distance $\boldsymbol{r} = |r_i - r_j|$ (Peebles 1980; Fisher et al. 1994). In the case of pairwise velocities for one of the high- or low-concentration (halo age) samples, the index i corresponds to galaxies in one of the quintiles of the secondary halo property, and the j index to the full galaxy sample corresponding to the chosen redshift and space density cut. As mentioned earlier, we will mostly focus on concentration as the second parameter. Results obtained using age will be shown for some interesting cases in the appendix.

Fig. 2 shows the pairwise velocities for selected samples both in their original and in their shuffled forms (in solid and dashed lines). Negative values correspond to pairs of galaxies that are moving towards one another. The case where pairwise velocities are calculated between all galaxies is shown in black; blue and red are for the 20 per cent lowest and 20 per cent highest concentrations against the entire sample, respectively. Solid lines show the original samples; dashed lines represent the shuffled catalogues. The shaded regions show the error on the mean obtained from 10 jackknife resamplings.

The bottom subdivisions in each panel show the ratio of the high- and low-concentration samples to the full ones, with the same line style as the main panels. The ratios are intended to highlight whether halo assembly bias or occupancy variation is responsible for the difference in the pairwise velocity amplitude. If at large separations the ratio of shuffled samples is significantly different from unity, then the halo assembly bias is behind the amplitude differences. Only when occupancy variation is solely responsible will this ratio be equal to one.

The panels of the figure correspond to selections intended to represent the SDSS Main Galaxy sample (top left), the Emission Line Galaxies of eBOSS (top right), a $z = 1$ CMASS-like sample (bottom left), and a HETDEX-like survey (bottom right); the space densities and selection criteria are shown by the key in each panel. The bottom left-hand panel and the top right-hand panel correspond to the same redshift and space density but for different selections, stellar masses, and star formation rates, and highlight the importance of taking into account the physical property on which the selection is performed.

The pairwise velocity function shows a similar shape for the different samples, at least qualitatively. At small separations it starts off with a low-amplitude infall that increases towards the transition to 2-halo-term scales. This is mostly due to the mix of first-infall satellites with the ones that are in virial equilibrium within the halo; the fraction of first infalls increases with separation. At larger separations, $r \sim 3\,h^{-1}$ Mpc and above, infall velocities drop monotonically as the enclosed overdensities gradually become smaller. These results are qualitatively similar to the pairwise velocities of dark matter haloes reported by Xu & Zheng (2018).

In all cases, the galaxies residing in haloes in the lowest concentration quintile show the higher velocity infalls on the two-halo term, regardless of whether the selection is done using stellar mass or star formation rate. The amplitude of the difference is similar regardless of redshift, space density, or selection by either stellar mass or star formation rate. This trend extends down to the one-halo term, although it is noisier for the SFR selections, and for the lower density samples. However, since the selection is made using different galaxy properties, the average velocity values are different. In particular, the stellar mass selection shows a higher amplitude of infall.

At equal space density and redshift (bottom left-hand and top right-hand panels of Fig. 2) the amplitude of infall in the one-halo term is larger for the SFR selection, but the difference between low and high concentration is less significant than in the stellar mass selection case. This points to a larger spread in infall amplitudes when star-forming galaxies are selected. We do not pursue the reasons for this difference in this study.

We analyse the different panels of Fig. 2 to assess the role that halo assembly bias and occupancy variation have, on their own, on the amplitude of clustering when the halo concentration varies. As can be seen, the $z = 0$ case of a stellar-mass-selected sample illustrating the SDSS Legacy Sample shows pairwise velocities that differ by ∼20 per cent on two-halo-term scales for high and low concentrations, but almost no difference for the shuffled catalogues separated by concentration. This indicates that an important fraction of the effect comes from occupancy variation rather than halo assembly bias on its own. However, in the other panels, the shuffled and original catalogues present more similar behaviours, and these become almost the same for the $z = 2$ case. This points to a trend of decreasing importance of occupancy variation as the redshift increases. This is in agreement with Contreras et al. (2019), who report a similar trend of a smaller effect on the correlation function from occupancy variation as the redshift increases, for samples with low and high concentrations, for both SFR- and stellar-mass-selected samples.

The analysis of relative contributions of halo assembly bias and occupancy variation depends on the specific secondary halo property that is analysed. For samples selected by halo age, for the same cases shown in this figure, the results are qualitatively similar. For the SFR-selected samples, irrespective of redshift, the impact on the two-halo term is smaller than for the concentration cuts for both halo assembly bias and occupancy variation; these results are shown in Fig. A1 of the appendix.

Also, in the appendix, we redo this analysis using just the central galaxies, and find only small differences with respect to the ones presented in Fig. 2, of the order of 10 per cent. The results for the stellar mass and star formation rate samples at $z = 1$ and $n = 0.000316\,h^3\,\mathrm{Mpc}^{-3}$ can be seen in Fig. A2. This is consistent with Zehavi et al. (2018), who also find that the assembly bias in clustering is dominated by the central galaxy population.

4.2 Assembly bias and pairwise velocities

We also study the evolution of the effect of assembly bias on pairwise velocities. Fig. 3 shows the ratio of the average values of the pairwise velocities for the original and shuffled catalogues

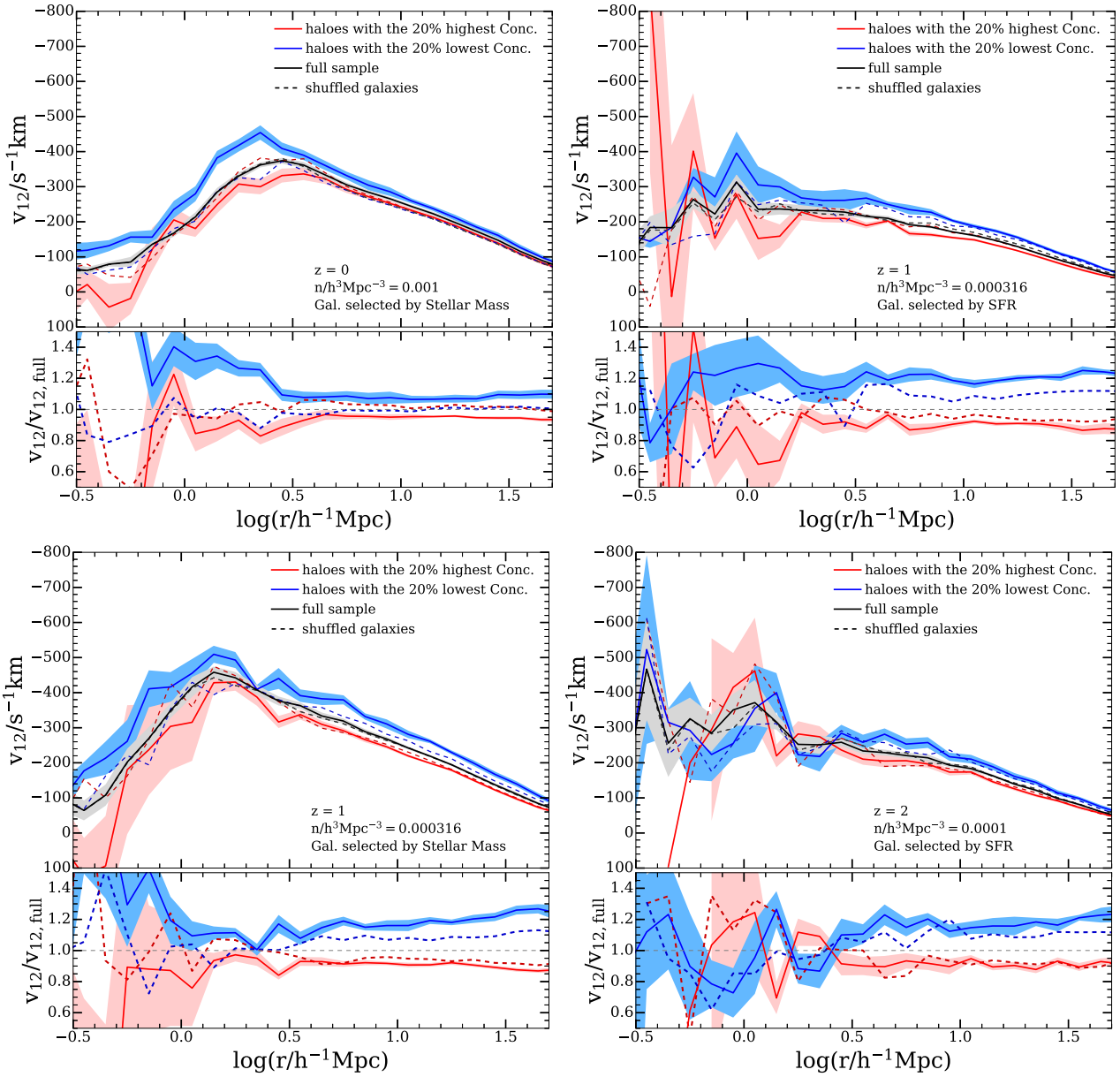

Figure 2. Pairwise velocities for all galaxies in samples selected by stellar mass for $z = 0$ and $z = 1$ (top and bottom left-hand panels, representing the SDSS main galaxies and a $z = 1$ CMASS-like sample, respectively), and galaxies selected by star formation rate for $z = 1$ and $z = 2$ (top and bottom right-hand panels, aimed at representing the SDSS-IV ELG and HETDEX samples, respectively), with space densities indicated in the key. The solid lines show the results for SAM galaxies in their original haloes, whereas the dashed lines are for the catalogues where galaxies were shuffled among haloes retaining the halo spatial and velocity structures. Black lines are for the full samples, red and blue for high and low concentrations, respectively.

as a function of redshift; note that here we are not dividing the sample into high and low values of the secondary parameter. When comparing between original and shuffled results, we are effectively looking at the fraction of the effect coming from occupancy variation since the shuffled samples are not affected by occupancy variation. The averaging is done on the two-halo-term range of separations $5 < r/h^{-1}\,\mathrm{Mpc} < 35$. By construction of the shuffled catalogues, this ratio is 1 in the one-halo-term regime. The top panel shows the results for stellar-mass-selected samples of different space densities, whereas the bottom panel shows this for SFR selections.

At low redshift, $z < 1$, for the stellar mass selection (top panel), the infall is stronger for the original catalogues irrespective of space density. This indicates that the previously reported effect of assembly bias that results in a stronger clustering for the unshuffled sample (Zehavi et al. 2018; Contreras et al. 2019) is also present for the relative velocities, which show a stronger infall prior to shuffling consistent with larger mass overdensities. As we move to higher redshifts we see that the importance of the occupancy variation becomes negligible, but as we reach $z > 2$, the infall becomes less pronounced in the unshuffled samples, a reversal of the behaviour at low redshift. This is also consistent with the reversal of the effect of occupancy variation seen in the clustering amplitude by Contreras et al. (2019). It should be noted that this reversal reflects the increase in the typical equivalent peak height of the sample with redshift; as early as Gao & White (2007) it was pointed out that at equal equivalent peak height, haloes of different concentrations show

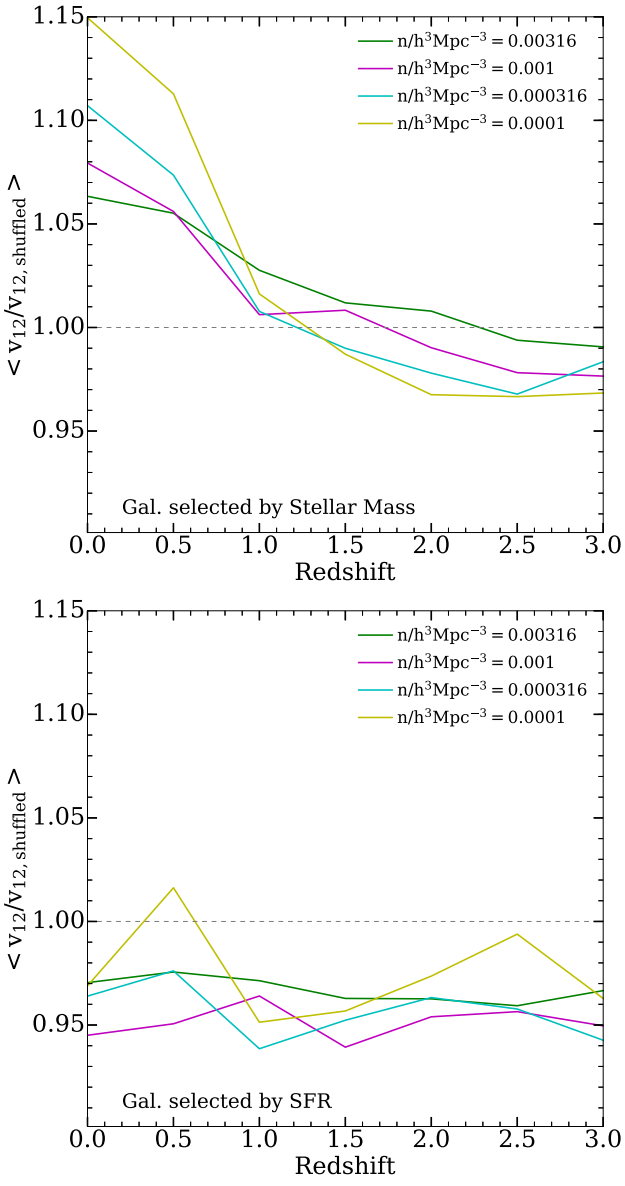

Figure 3. Ratio between averaged pairwise velocities over $5 < r/h^{-1}$ Mpc < 35 of the original and shuffled catalogues for stellar mass and SFR selections (top and bottom panels) for different galaxy number densities (different colours, indicated in the key).

consistent clustering amplitude differences regardless of redshift from $z = 0$ to $z = 3$. They noted that at low peak height values, clustering amplitude is correlated with concentration, whereas for high peak height values this is reversed, consistent with both our results and those of Contreras et al. (2019).

The bottom panel shows the same calculation for SFR-selected samples. As can be seen, there is no clear dependence on number density or redshift, and there is a tendency to exhibit weaker large-scale infall in the original samples. This points in the same direction as the results of Contreras et al. (2019) where the shuffled samples show more clustering than the unshuffled ones for SFR-selected samples.

The fact that the changes due to assembly bias in the infall velocities are consistent with those reported for the clustering suggests that the reason behind both is the change in the bias factor of the samples, as also pointed out by Xu & Zheng (2018). Later on in this work we will quantify whether the velocity differences can be explained by the change in the amplitude of clustering alone, i.e. by the change in the bias factor.

5 ANALYSIS IN REDSHIFT SPACE

Before delving into the quantitative analysis of the multipoles of the correlation function in redshift space, we will make a qualitative assessment of redshift-space distortions for samples of equal halo mass but different secondary properties.

5.1 Redshift-space distortions of the correlation function

We measure the correlation functions as a function of projected and line-of-sight separations, $\xi(r_p, \pi)$, using the positions of galaxies in the simulation with the *z-axis* position distorted by the peculiar velocities as mentioned above. We calculate autocorrelations for our samples of different space densities and redshifts, and the corresponding cross-correlation functions between these full samples and quintiles with high or low concentration. This allows us to see the effect of the secondary halo parameter, i.e. assembly bias, on the redshift-space correlation functions.

Fig. 4 shows the actual redshift-space distortion pattern for low and high halo concentrations (left- and right-hand panels, respectively) for stellar-mass- and SFR-selected samples (top and bottom) at $z = 1$, with equal space densities. This figure shows the typical pattern of redshift-space distortions, where at small projected separations the equal correlation function amplitude contours are elongated in the direction of the line of sight due to the random motions inside virialized structures (the so called fingers of God effect). At larger projected separations, the contours become squashed in the line of sight due to the coherent large-scale infall motions.

As can be seen the elongations at small r_p are present in all cases, but there are clear differences in both the amplitude and redshift-space distortion of the correlation functions between high- and low-concentration samples, where the latter show a less prominent infall pattern (squashing of contours). There are also hints of a stronger infall pattern in the samples selected by the galaxy star formation rate but this is probably due to a combination of smaller fingers of God effect and the actual infall, because the amplitude of the correlation function is higher for the stellar mass selection, as is also the case for the amplitude of the pairwise velocities (see Fig. 2) . Left-hand panels with the low-concentration samples show a higher clustering amplitude.

5.2 Multipoles

We quantify the effect of redshift-space distortions on the correlation function by measuring the multipoles $\xi_l(s)$,

$$\xi_l(s) = \frac{2l+1}{2} \int \mathrm{d}\mu \xi(r_\mathrm{p}, \pi) L_l(\mu), \quad (3)$$

where $s = \sqrt{r_\mathrm{p}^2 + \pi^2}$ is the separation in redshift space, l is the multipole moment, $\mu = \pi/s$, and $L_l(\mu)$ are Legendre polynomials of order l (Hamilton 1992; Cole et al. 1994; Padmanabhan & White 2008). Fig. 5 shows the monopole, quadrupole, and hexadecapole ($l = 0$, 2, 4, respectively) for the samples of galaxies selected by stellar mass and star formation rate (top and bottom panels) at $z = 1$ for a space density of $n = 0.000316\,h^3\,\mathrm{Mpc}^{-3}$. Solid lines

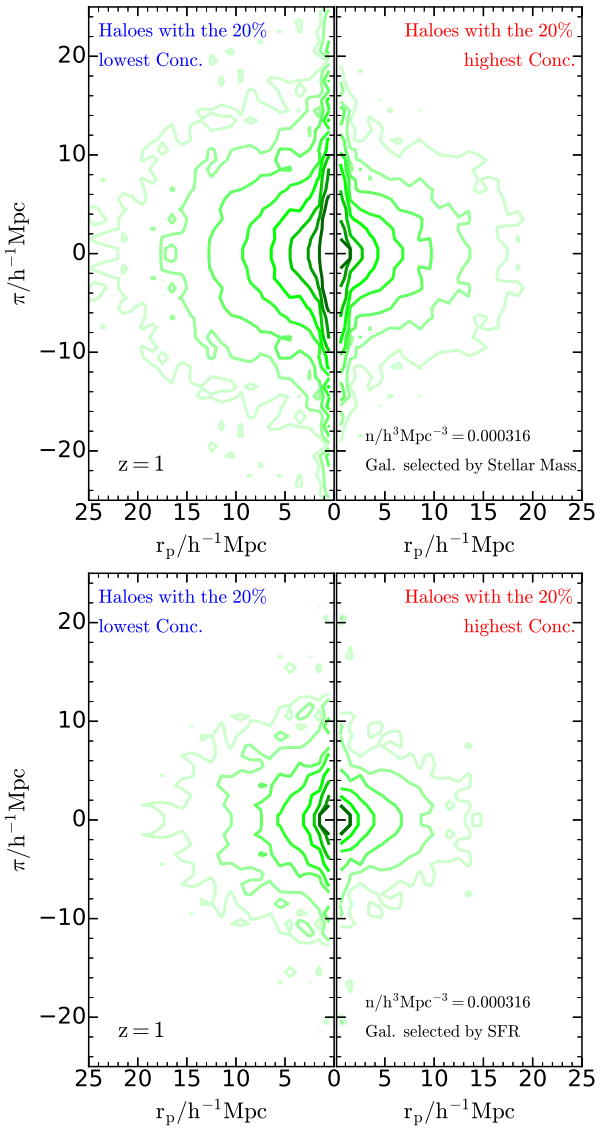

Figure 4. Correlation functions as a function of separation in the directions parallel (π) and perpendicular to the line of sight (r_p), in this case coincident with the z-axis of the simulation box. Both panels correspond to $z = 1$ and a space density of $n = 0.000316\,h^3$ Mpc^{-3}, for stellar-mass- and SFR-selected samples (top and bottom, respectively), for low and high concentrations (left and right, respectively). The lines show contours of equal correlation function amplitudes; darker shades are for higher amplitudes of $\log \xi = -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1$, and 1.25.

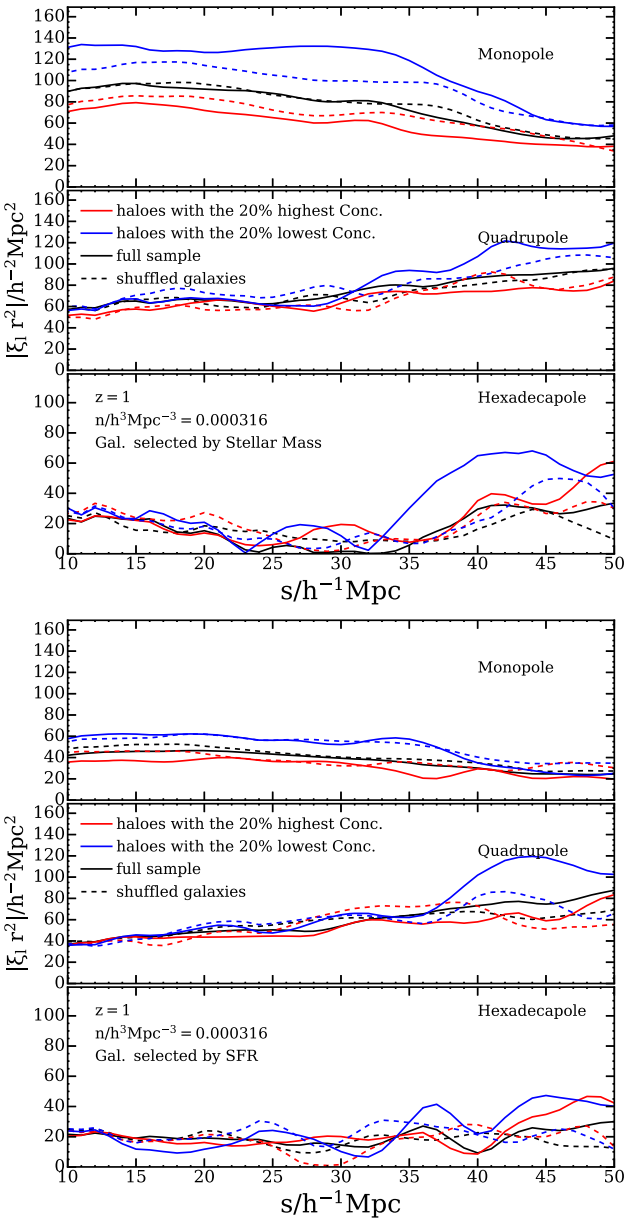

Figure 5. Multipoles of the cross-correlation function for $z = 1$ and a space density of $n = 0.000316\,h^3$ Mpc^{-3}, for stellar mass (top) and star formation rate (bottom) selected samples. Each panel is further subdivided to show the monopole, quadrupole, and hexadecapole ($l = 0, 1, 2$) as a function of separation in redshift space s. Black lines show the full sample, blue and red the first and fifth quintiles in concentration, and solid and dashed lines the original and shuffled samples, respectively.

correspond to the original samples, and dashed lines to the shuffled samples.

Irrespective of whether star formation rate or stellar mass is used to select the samples, we find that the monopole shows results that are consistent with the real-space correlations reported by Contreras et al. (2019) for both assembly bias and occupancy variation. Namely that the low concentration samples show stronger clustering. This was not necessarily expected since the redshift-space monopole responds in some degree to redshift-space distortions which are also influenced by assembly bias and occupancy variation, as shown in the previous section. This effect is stronger for the stellar mass selection. The shuffled catalogues continue to show this trend but to a lesser degree; the small differences with respect to the full samples show that there is an effect from assembly bias.

The quadrupole is entirely due to the effect of redshift-space distortions. For both stellar mass and SFR selections the quadrupole shows a higher amplitude for the low-concentration samples, particularly for the largest scales r/h^{-1} Mpc > 30. The resulting quadrupoles for the shuffled samples appear qualitatively similar but with a lower amplitude, pointing to an effect coming mostly from halo assembly bias as in the shuffled samples there is no occupancy variation. The hexadecapoles are consistent between samples with low and high concentration, and for the shuffled

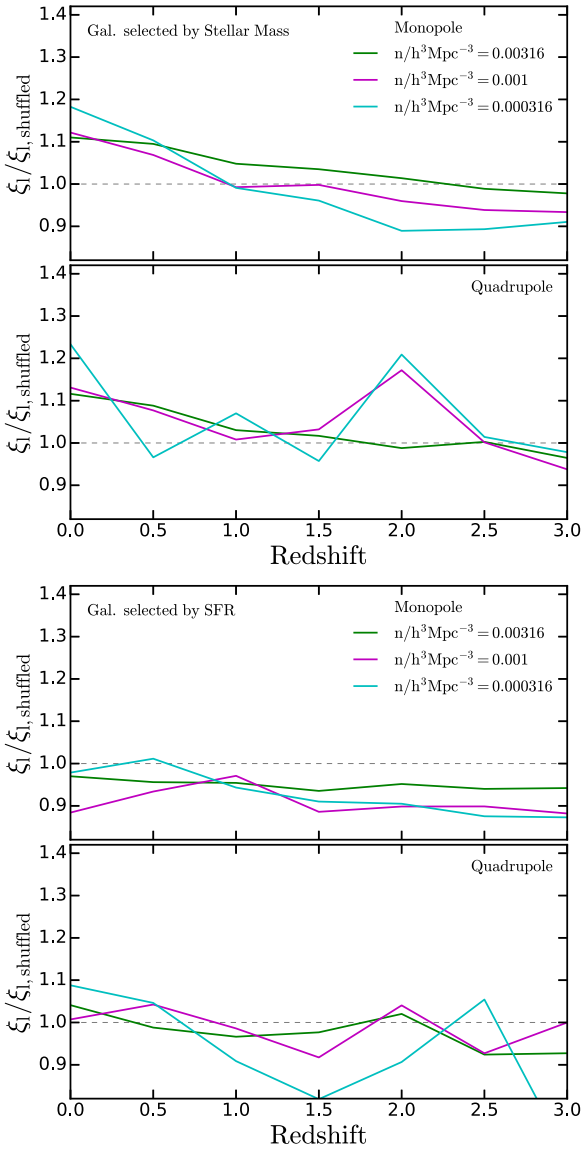

Figure 6. Ratio between averaged multipole moments of the redshift-space correlation functions for $5 < r/h^{-1}\,\mathrm{Mpc} < 35$; this range isolates the scales on which the two-halo term is important. The ratios are shown as a function of redshift for different space densities (colours as indicated in the key) and for stellar mass (top) and star formation rate selections (bottom). The top and bottom subdivisions of panels correspond to the monopole and quadrupole, respectively.

and unshuffled catalogues, at least to within the uncertainties of our measurements. By averaging the multipoles over two-halo-term scales these differences can be studied in more detail.

Fig. 6 shows the average values of the ratios between multipoles for the original and shuffled results. We calculate the average over the range $5 < r/h^{-1}\,\mathrm{Mpc} < 35$. The top panels show the ratio for stellar-mass-selected samples for three representative space densities. As can be seen, the original samples show a stronger monopole at low redshifts and this difference decreases for higher redshifts. By $z = 1$ the difference becomes less significant, and for the two lowest space densities shown in the figure the difference tends to reverse, making the monopole stronger for the shuffled samples. The latter suggests that occupancy variation tends to shift from preferentially weighting high-mass haloes at low redshifts to low-mass ones at $z > 1$, thereby lowering the clustering amplitude. This is similar to the behaviour of the infall strength as shown in Fig. 3. Furthermore, the ratios for the SFR-selected samples of the monopole for the original and shuffled samples are also consistent with the results for the infall velocities.

On two-halo-term scales, the quadrupole is expected to show a direct relation with the infall velocities, as the latter produce the flattening of the iso-correlation function contours that are quantified by the quadrupole. As a result, we find similar trends to those shown in Fig. 3 and to the results for the monopole. Namely, the effects from occupancy variation on clustering and infall velocities are consistent with one another, at least qualitatively. Even though the quadrupole showed stronger differences at $r/h^{-1}\,\mathrm{Mpc} > 30$ (Fig. 5), using this range of separations for the quadrupole does not influence our conclusions on the average ratios.

6 COSMOLOGICAL PARAMETERS: β

In the previous sections we showed the effects of assembly bias and occupancy variation on the pairwise infall velocities and redshift-space clustering of galaxies. We found that in general there is a qualitative agreement between the results for velocities and clustering in the sense that samples that show stronger clustering also show stronger pairwise velocities as expected if bias is the main driver of both effects.

In this section we make a quantitative comparison between the two by studying the resulting values of the β parameter. This parameter represents the ratio between the logarithmic growth rate $f(\Omega)$, which depends on the matter density parameter, and the bias parameter of a sample,

$$\beta = \frac{f(\Omega)}{b_s}. \tag{4}$$

The β parameter can be obtained from the ratio between the monopoles of the correlation functions in real and redshift space (Kaiser 1986),

$$\xi_0(s) \equiv \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right)\xi(r). \tag{5}$$

Since we are measuring cross-correlations, b_s from equation (4) is the bias of the sample with high or low halo concentration (Hamilton 1992; Peacock 1999). We show the values of this parameter in Fig. 7 for a galaxy sample with $n = 0.000316\,h^3\,\mathrm{Mpc}^{-3}$ selected by their stellar mass at $z = 1$ (illustrative of a $z = 1$ CMASS eBOSS-like galaxy sample). We also measure β using multipoles as is done with real data; the results are shown in the appendix (Fig. A3). We do not compare to the expected value of β because rather than being interested in the ability of this particular redshift-space distortion model to reproduce the measured value of this parameter, we are interested in how its value changes due to halo assembly bias and occupancy variation. Different models for redshift-space distortions (RSD) will be tested in Jimenez et al. (in preparation).

For this sample the ratio of the monopoles is lower for galaxies in low-concentration haloes; note that this difference can be due to the different bias factors of low- and high-concentration haloes. We also see that β is roughly constant between separations of 15 and $35\,h^{-1}\,\mathrm{Mpc}$ as expected from linear theory. The dashed lines show the resulting values of β for the shuffled catalogues (full, low, and high concentration). As can be seen the results for high and low concentration are similar for the original and shuffled samples (little effect from occupancy variation). The differences between high and low concentration are consistent with the pairwise velocity results,

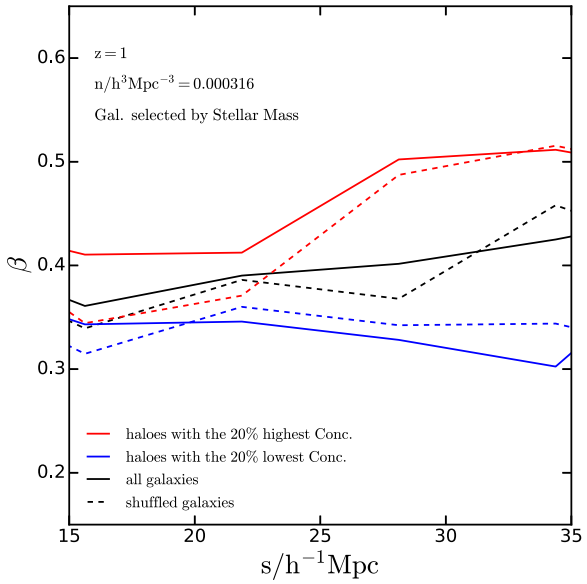

Figure 7. Values of the β parameter as a function of separation for a stellar-mass-selected sample with $n = 0.000316\,h^3\,\mathrm{Mpc}^{-3}$ (solid black line), and for the quintiles of high and low halo concentration (solid red and blue lines, respectively), as a function of separation. The dashed lines are for the corresponding shuffled samples.

and also with Contreras et al. (2019), giving an indication that the change in β is driven by the change in the bias factor of the sample.

Fig. 8 shows the averaged ratio of β for the high- (β_{highC}) and low-concentration selection (β_{lowC}), over the range $s = 15$–$35\,h^{-1}$ Mpc for different number density samples selected by cuts in stellar mass (see the figure key) as a function of redshift. As can be seen, this ratio evolves with redshift possibly due to bias evolution as we evaluate below. Its dependence is qualitatively similar for the original and shuffled samples. At $z = 0$ the β values for low and high concentrations are consistent with one another but by $z \sim 1$ the high-concentration samples show a higher β parameter by about 20 per cent. At $z = 3$ in both the original and the shuffled samples, the ratio is again consistent with $\beta_{\mathrm{highC}}/\beta_{\mathrm{lowC}} = 1$. Errorbars show the scatter in the ratios from jackknife resamplings.

The figure also shows as dashed lines the square root of the inverse ratio between the cross-correlation functions of the same samples, averaged over the same range of separations. This quantity corresponds to the ratio of bias factors of the samples, i.e. the denominator of equation (4). As the dashed lines are consistent with the solid ones, we can conclude that galaxy assembly bias does not introduce any detectable systematics on estimates of $f(\Omega)$. We have also repeated this analysis for SFR-selected samples (see Fig. A4 in the appendix), reaching the same conclusion.

As galaxy assembly bias includes both the effect of halo assembly bias and occupancy variation, we also investigate whether halo assembly bias alone produces any systematic effects on β in the bottom panel of Fig. 8, which shows the result for the shuffled samples. No systematic effects can be seen in these results either to the level of precision of our analysis (note that this is lower than in Chen et al. 2018). This conclusion should also hold for more complex RSD modelling since according to these results assembly bias and occupancy variation simply change the bias of the sample of galaxies, and it is possible to use this value to obtain accurate estimates of total mass overdensities. We will further investigate this in Jimenez et al. (in preparation).

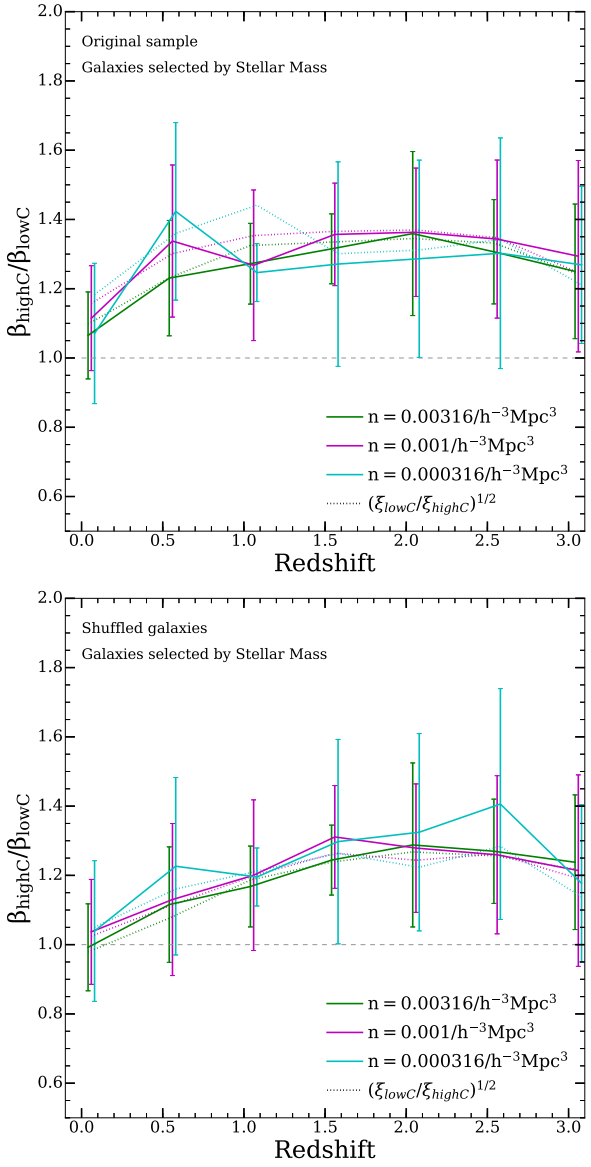

Figure 8. Ratio between the averaged β values for high- and low-concentration samples, as a function of redshift, for different space density samples (colours, as indicated in the key) selected by their stellar mass. Solid lines show results for the original samples; dashed lines show the square root of the ratio between the real-space correlation functions for the same samples, averaged over the same range of separations; and the agreement between solid and dashed lines shows that both velocities and clustering respond to assembly bias in similar ways. The horizontal grey dashed line shows the unit ratio. The top panel corresponds to the samples as directly obtained from the simulation, the bottom panel to the shuffled ones. Errors are obtained from jackknife resamplings.

7 CONCLUSIONS

We studied how galaxy assembly bias, the phenomenon by which the clustering of galaxies depends not only on the halo occupation distribution and halo mass, but also on secondary halo properties, affects peculiar velocity statistics and their relation to the spatial distribution of galaxies. We use the G13 semi-analytic model run on the Millenium *WMAP*-7 *N*-body simulation to select samples of galaxies using stellar mass and star formation rate cuts, such that the space density of the samples spans several orders of magnitude,

with redshifts ranging from $z = 0$ to $z = 3$. We pay special attention to those that more closely mimic the SDSS main galaxy sample, a CMASS-like and the ELG samples at $z = 1$, and $z = 2$ HETDEX galaxies. This work is the third in a series studying galaxy assembly bias in SAMs following Zehavi et al. (2018) and Contreras et al. (2019), where we move our focus from spatial clustering to velocity space.

Our work complements Xu & Zheng (2018); while they concentrate on assembly bias of haloes, looking at several secondary properties, and how these produce correlated changes in pairwise velocities and clustering of haloes, we focus on semi-analytic galaxies, and study samples of equal halo mass and different concentrations. We also test whether cosmological parameter constraints using redshift space distortions could be affected by systematics when not taking into account the assembly bias effect.

In general we find that the peculiar velocity field shows the expected response to halo assembly bias and occupancy variation; galaxy velocities respond to the bias factor of the sample regardless of whether the halo mass or other halo property is driving the change in the bias factor of the sample. We summarize our results as follows.

(i) Low-redshift, high-space-density, stellar-mass-selected samples such as the SDSS Main Galaxy Sample show strong occupancy variation effects on both the spatial correlations and pairwise velocities. On the other hand, halo assembly bias gains importance for both stellar-mass- and SFR-selected samples at $z = 1$ such as CMASS-like and ELGs. At higher redshifts, the occupancy variation appears to have no effect on the clustering and velocity variations, with most of the effect coming from halo assembly bias, at least for the sparse sample that will be available at $z \sim 2$ from surveys such as HETDEX.

(ii) The amplitude of the infall is affected by occupancy variation in a similar way to what is seen in the spatial clustering (Contreras et al. 2019); the shuffled catalogue shows a smaller difference of large-scale infall amplitude with concentration.

(iii) The results for the multipoles of the redshift-space correlation function show that both halo assembly bias and occupancy variation affect the monopole and the quadrupole. The size of the simulation does not allow us to reach firm conclusions regarding the effect on the hexadecapole. We use the monopole and the real-space correlation function to estimate the β parameter, and find that its changes are driven fully by the variation of the bias parameter with halo mass or secondary halo property.

To the level of accuracy allowed by the simulation, and to the level of the physics taken into account by a semi-analytic model, we find that assembly bias does not introduce further systematics to cosmological parameter estimates from redshift-space correlation functions in samples similar to the SDSS main galaxy sample, $z = 1$ CMASS-like or ELG eBOSS, or HETDEX samples.

Our results show that halo assembly bias, in particular, affects both clustering and velocities in a consistent way. This is yet another indication that halo mass alone is too simple a proxy for the equivalent peak height of the original overdensities that evolved into haloes via complex non-linear physics.

ACKNOWLEDGEMENTS

We thank Raul Angulo, Alvaro Orsi, and the referees David Weinberg and Andres Salcedo, for helpful comments and discussions. This work was made possible by the efforts of Gerard Lemson and colleagues at the German Astronomical Virtual Observatory in setting up the Millennium Simulation data base in Garching. NP and SC acknowledge support from a Science and Technology Facilities Council STFC/Newton-CONICYT Fund award (ST/M007995/1 - DPI20140114) and Anillo ACT-1417. NP is further supported by 'Centro de Astronomía y Tecnologías Afines' BASAL AFB-170002 and byFondecyt Regular 1150300. SC acknowledges support of the European Research Council through grant number ERC-StG/716151. IZ acknowledges support by National Science Foundation grant AST-1612085. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement No. 734374. The calculations for this paper were performed on the ICC Cosmology Machine, which is part of the DiRAC-2 Facility jointly funded by STFC, the Large Facilities Capital Fund of BIS, and Durham University and on the Geryon computer at the Center for Astro-Engineering UC, part of the BASAL PFB-06, which received additional funding from QUIMAL 130008 and Fondequip AIC-57 for upgrades.

REFERENCES

Adams J. J. et al., 2011, ApJS, 192, 5
Artale M. C., Zehavi I., Contreras S., Norberg P., 2018, MNRAS, 480, 3978
Biviano A. et al., 2017, A&A, 607, A81
Bullock J. S., Dekel A., Kolatt T. S., Kravtsov A. V., Klypin A. A., Porciani C., Primack J. R., 2001, ApJ, 555, 240
Busch P., White S. D. M., 2017, MNRAS, 470, 4767
Campbell D., van den Bosch F. C., Hearin A., Padmanabhan N., Berlind A., Mo H. J., Tinker J., Yang X., 2015, MNRAS, 452, 444
Chen J., Zhang P., Zheng Y., Yu Y., Jing Y., 2018, ApJ, 861, 58
Chuang C.-H., Wang Y., 2012, MNRAS, 426, 226
Cole S., Aragon-Salamanca A., Frenk C. S., Navarro J. F., Zepf S. E., 1994, MNRAS, 271, 781
Comparat J. et al., 2016, A&A, 592, A121
Contreras S., Zehavi I., Baugh C. M., Padilla N., Norberg P., 2017a, MNRAS, 465, 2833
Contreras S., Padilla N., Lagos C. D. P., 2017b, MNRAS, 472, 4992
Contreras S., Zehavi I., Padilla N., Baugh C. M., Jiménez E., Lacerna I., 2019, MNRAS, 484, 1133
Croton D. J. et al., 2006, MNRAS, 365, 11
Croton D. J., Gao L., White S. D. M., 2007, MNRAS, 374, 1303
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Dawson K. S. et al., 2016, AJ, 151, 44
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
De Lucia G., Kauffmann G., White S. D. M., 2004, MNRAS, 349, 1101
Eisenstein D. J. et al., 2011, AJ, 142, 72
Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J. P., 1994, MNRAS, 267, 927
Forero-Romero J. E., Contreras S., Padilla N., 2014, MNRAS, 443, 1090
Gao L., White S. D. M., 2007, MNRAS, 377, L5
Gao L., White S. D. M., Jenkins A., Stoehr F., Springel V., 2004, MNRAS, 355, 819
Gao L., Springel V., White S. D. M., 2005, MNRAS, 363, L66
Guo Q. et al., 2011, MNRAS, 413, 101
Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, MNRAS, 428, 1351 (G13)
Guo H. et al., 2015, MNRAS, 446, 578
Guo Q. et al., 2016, MNRAS, 461, 3457
Guzzo L. et al., 2008, Nature, 451, 541
Hamilton A. J. S., 1992, ApJ, 385, L5
Hearin A. P., 2015, MNRAS, 451, L45
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, MNRAS, 431, 3373
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2663
Hu W., Haiman Z., 2003, Phys. Rev. D, 68, 063004
Kaiser N., 1986, MNRAS, 222, 323

Kazin E. A., Sánchez A. G., Blanton M. R., 2012, MNRAS, 419, 3223
Lacerna I., Contreras S., González R. E., Padilla N., Gonzalez-Perez V., 2018, MNRAS, 475, 1177
Leja J. et al., 2013, ApJ, 778, L24
Mao Y.-Y., Zentner A. R., Wechsler R. H., 2018, MNRAS, 474, 5143
McEwen J. E., Weinberg D. H., 2018, MNRAS, 477, 4348
Mundy C. J., Conselice C. J., Ownsworth J. R., 2015, MNRAS, 450, 3696
Okumura T., Matsubara T., Eisenstein D. J., Kayo I., Hikage C., Szalay A. S., Schneider D. P., 2008, ApJ, 676, 889
Padilla N. D., Christlein D., Gawiser E., González R. E., Guaita L., Infante L., 2010, MNRAS, 409, 184
Padmanabhan N., White M., 2008, Phys. Rev. D, 77, 123540
Peacock J. A., , 1999, Cosmological Physics. Cambridge Univ. Press, Cambridge
Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
Salcedo A. N., Maller A. H., Berlind A. A., Sinha M., McBride C. K., Behroozi P. S., Wechsler R. H., Weinberg D. H., 2018, MNRAS, 475, 4411
Sato-Polito G., Montero-Dorta A. D., Abramo L. R., Prada F., Klypin A., 2018, preprint (arXiv:1810.02375)
Seo H.-J., Eisenstein D. J., Zehavi I., 2008, ApJ, 681, 998
Sheth R. K., Tormen G., 2004, MNRAS, 350, 1385
Shoji M., Jeong D., Komatsu E., 2009, ApJ, 693, 1404
Sin L. P. T., Lilly S. J., Henriques B. M. B., 2017, MNRAS, 471, 1192
Springel V. et al., 2005, Nature, 435, 629
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Strauss M. A. et al., 2002, AJ, 124, 1810
Tinker J. L., Hahn C., Mao Y.-Y., Wetzel A. R., Conroy C., 2018, MNRAS, 477, 935
Torrey P. et al., 2015, MNRAS, 454, 2770
Wagner C., Müller V., Steinmetz M., 2008, A&A, 487, 63
Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, ApJ, 652, 71
Xu X., Zheng Z., 2018, MNRAS, 479, 1579
Ye J.-N., Guo H., Zheng Z., Zehavi I., 2017, ApJ, 841, 45
Zehavi I., Contreras S., Padilla N., Smith N. J., Baugh C. M., Norberg P., 2018, ApJ, 853, 84
Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, MNRAS, 443, 3044
Zu Y., Mandelbaum R., 2016, MNRAS, 457, 4360
Zu Y., Zheng Z., Zhu G., Jing Y. P., 2008, ApJ, 686, 41

APPENDIX A: ADDITIONAL RESULTS

In this appendix we present several additional results that are mentioned in the main text but which we felt would have distracted the reader from the flow of the paper.

In the main text we concentrated our analysis on halo concentration as the secondary halo parameter. Fig. A1 shows the results for the pairwise velocities when using halo age as the secondary parameter. In this case the results indicate that halo assembly bias and occupancy variation roughly cancel one another in the $z = 0$ sample, as evidenced by a noticeable difference in the pairwise velocities of early- and late-formed haloes for the shuffled samples, with younger haloes showing a lower amplitude of infall velocity. In this case halo assembly bias increases the infall for early-formed haloes, and occupancy variation lowers it for the same population.

Differences in pairwise velocities are also present in the $z = 1$ SFR-selected samples of the top right-hand panel. At this redshift there is a net galaxy assembly bias effect on velocities, and again halo assembly bias imprints the opposite trend than occupancy variation. At the same redshift, samples selected by stellar mass show only an occupancy variation effect (lower left). The $z = 2$ example on the lower right shows effects from occupancy variation alone, although small, and no halo assembly bias. These results are consistent with those found by Contreras et al. (2019) for the clustering amplitudes of samples separated by halo age as secondary property at different redshifts.

We show the pairwise velocities for different quintiles of concentration for central galaxies only in Fig. A2. By comparing to Fig. 2 it is clear that the results remain unchanged, with similar amplitude differences for high- and low-concentration samples selected by either stellar mass or SFR at $z = 1$. The differences between shuffled samples are also similar to when using the full population of galaxies. This indicates that centrals dominate the amplitude of the pairwise velocities in our chosen samples, and that the shuffled catalogues simply remove the dependence of the minimum halo mass to host a central galaxy on secondary halo parameters. These results are consistent with Zehavi et al. (2018).

In the main text, we calculated the β parameter using the ratio of the monopoles of the correlation functions in redshift and real space (see Fig. 7). Fig. A3 shows the results from combining equations (26) and (27) of Hamilton (1992), which is the method that can be adopted when analysing actual observational data that consist of galaxies with distances measured via their redshifts. As can be seen the results are roughly consistent with those from the monopoles, with galaxies in the original sample with higher concentration showing a larger value of β at all scales. Qualitatively this trend is also present in the shuffled samples, but the amplitude difference is smaller; these differences should be revisited with a larger simulation in future work.

Fig. A4 shows the ratio between β values for samples in the upper and lower concentration quintiles in SFR-selected samples. This confirms the results already obtained for stellar-mass-selected samples (Fig. 8), that the change in the β parameter obtained from redshift-space distortions for samples in the extreme quintiles of halo concentration simply responds to their change in bias factor introduced by the split in concentration. The figure is for the original samples; the shuffled samples show equivalent results.

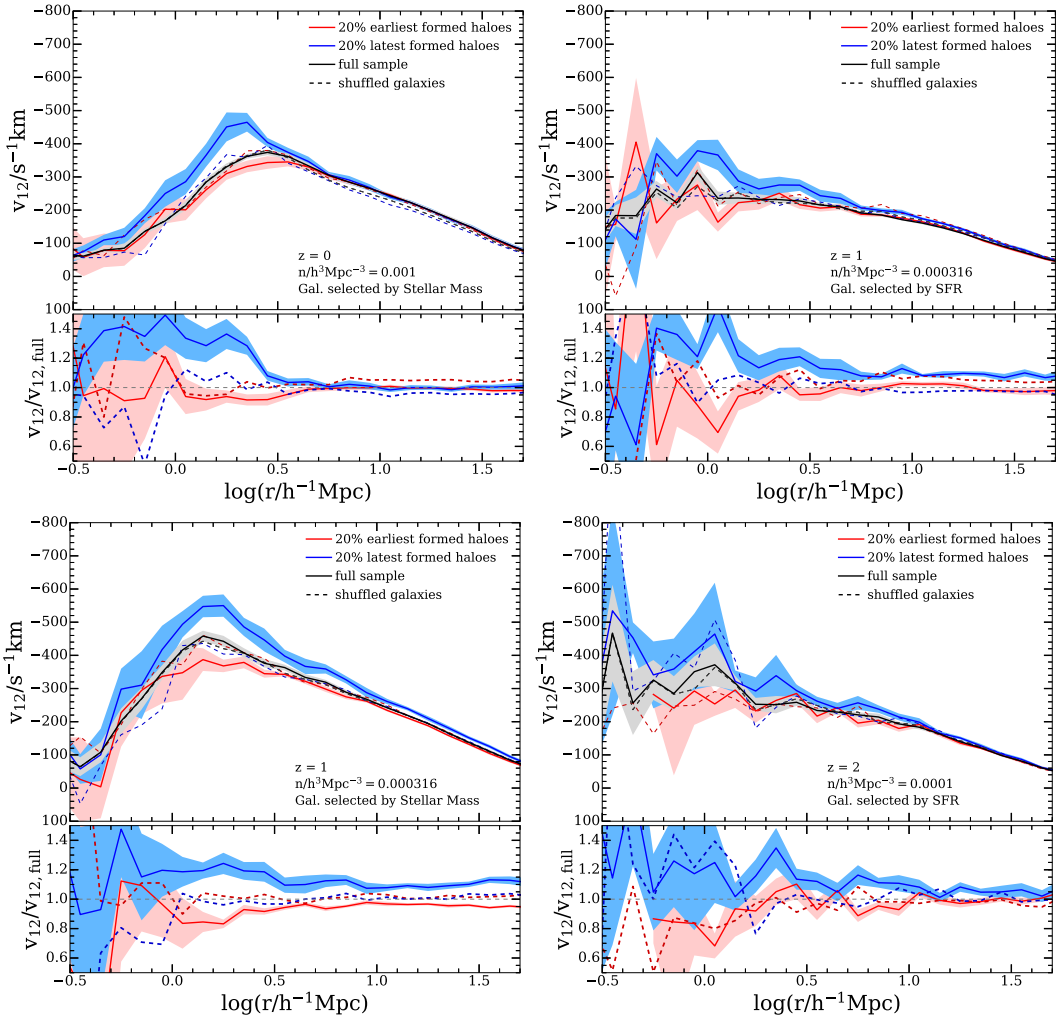

Figure A1. Pairwise velocities for all galaxies in samples selected by stellar mass for $z = 0$ and $z = 1$ (top and bottom left-hand panels, respectively), and galaxies selected by star formation rate for $z = 1$ and $z = 2$ (top and bottom right-hand panels, respectively), with space densities indicated in the key. The blue and red lines correspond to the first and fifth quintiles in halo age. The solid lines show the results for SAM galaxies in their original haloes, whereas the dashed lines are for the catalogues where galaxies were shuffled among haloes retaining the halo spatial and velocity structures.

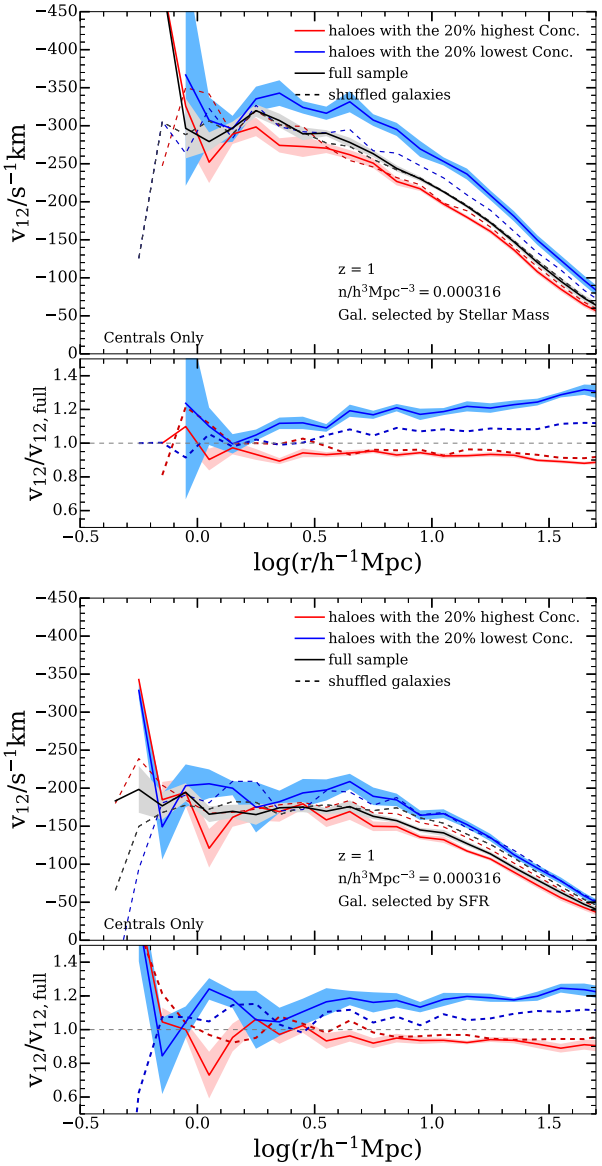

Figure A2. Pairwise velocities for central galaxies in the samples for stellar mass (top) and star formation rate selection (bottom) for redshift $z = 1$ and 2 (top and bottom, respectively). The case where only central galaxies are used shows a similar result for the two-halo term, pointing to an assembly bias effect dominated by central pairs.

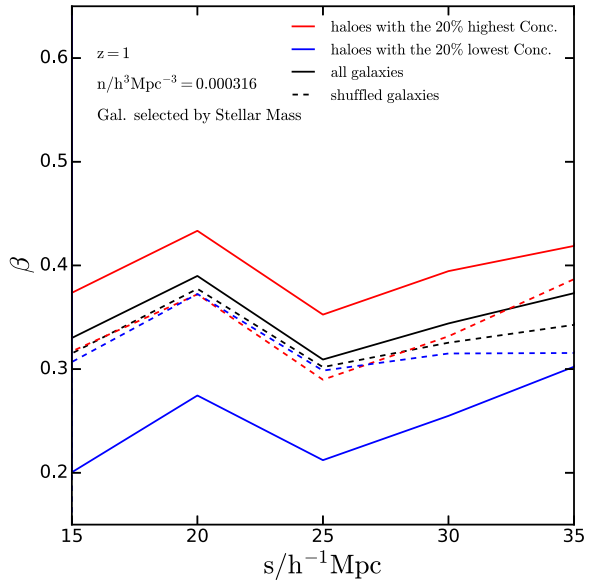

Figure A3. Values of the β parameter as a function of separation obtained from the multipoles of the correlation function in redshift space. The results shown correspond to a stellar-mass-selected sample with $n = 0.000316\,h^3\,\mathrm{Mpc}^{-3}$ (solid black line), and for the quintiles of high and low halo concentration (solid red and blue lines, respectively), as a function of separation. The dashed lines are for the corresponding shuffled samples.

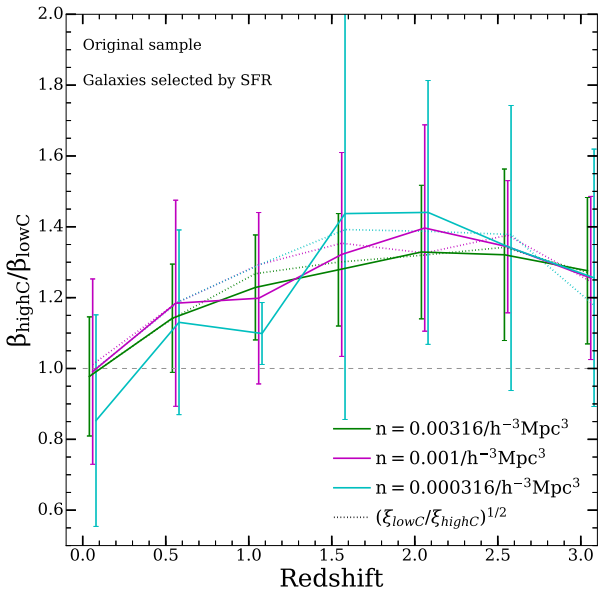

Figure A4. Ratio between the averaged β values for high- and low-concentration samples, as a function of redshift, for different space density samples (colours, as indicated in the key) selected by their star formation rate. Solid lines show results for the original samples; dashed lines show the square root of the ratio between the real-space correlation functions for the same samples, averaged over the same range of separations; the agreement between solid and dashed lines shows that both velocities and clustering respond to assembly bias in similar ways. The horizontal grey dashed line shows the unit ratio. Errors are obtained from jackknife samples.

This paper has been typeset from a TEX/LATEX file prepared by the author.